JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 96, 24–41 (1983)

Probability on Fuzzy Sets and ϵ-Fuzzy Sets

J. F. BUONCRISTIANI

Mathematics Department, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Submitted by G.-C. Rota

This research investigates the development of probability calculi for fuzzy sets (abbreviated FS's). Two FS algebras are employed, each as a basis for a FS probability calculus. One employs minima for FS intersections and consequently maxima for unions. The other uses a product rule for intersections and a standard probability-like addition rule for unions. The two derived FS probability theories are not identical and neither is identical with standard probability theory. Finally, ε-fuzzy sets, or almost standard sets, are defined and investigated. The main result of the last section is that while FS's do not satisfy all the laws of Boolean algebra fuzzy sets that are almost standard almost satisfy these laws. Moreover, the extent to which such laws are satisfied decreases as their complexity increases.

INTRODUCTION

In 1965 Zadeh [6] defined the notion of a fuzzy set (abbreviated FS) in order to represent inexact concepts or vague precidates. The probability of a FS or fuzzy event was later defined by Zadeh in 1968 [7]. Since then there have been reports on the development of a theory of possibility using FS's [3, 4, 8] but no thorough, systematic inquiry into the fundamental notions of standard probability theory from a FS point of view. This is part of the goal of this research paper.

We begin by using two FS algebras in order to develop two FS probability calculi. The two FS algebra chosen here are simple and current in the FS literature. The first algebra uses minima of FS membership functions for FS intersection and maxima for unions (M & M algebra). The second uses a product rule for intersections and a standard probability-like addition rule for unions (PR algebra). The FS probability theory (FSPT) developments along both algebraic directions will be interwoven. Different probability consequences will be obtained with the two FS algebras. Some probability theory statements are common to both and standard probability theory (SPT). But both FSPT's will be seen to be different from SPT.

In the first section of the paper preliminaries are taken care of. The notions of independence and conditional probability of FS's are investigated

0022-247X/83 $3.00
Copyright © 1983 by Academic Press, Inc.
All rights of reproduction in any form reserved.

in Section II. It will be demonstrated that with either algebra it may happen that a FS and its complement are independent and such that both have non-zero probability. With the M & M algebra independence of two FS's need not imply that of their complements but not so with the PR algebra. Fuzzy partitions and Bayes' Theorem are examined in Section III. The traditional form of Bayes' Theorem will be seen to hold in the PR case but not in the M & M case. In Section V the notion of not quite so fuzzy or an ε-fuzzy set (ε-FS) is defined and investigated. ε-FS's are almost standard sets so to speak. The main result in Section VI, which generalizes the developments of Section V, is a theorem which in effect states that while FS's do not satisfy all the laws of Boolean algebra, FS's that are almost standard almost satisfy these laws. Moreover, the extent to which such laws are satisfied decreases as their complexity increases.

I. Preliminaries: Probability of FS's

We adopt the conventional FS representation by its characteristic function. Let I denote the closed unit interval and Ω a background space, then a FS A is defined by its fuzzy characteristic function $\mu_A: \Omega \to I$ [2, 6]. The complement of A, A^c is defined by $\mu_{A^c}(\omega) = 1 - \mu_A(\omega)$ [2, 6]. We define FS intersection so that when De Morgan's law is applied the usual relationship on characteristic functions holds. Specifically, for FS's A and B

$$\begin{aligned} \mu_{A \cup B}(\omega) &= 1 - \mu_{A^c \cap B^c}(\omega) \\ &= \mu_A(\omega) + \mu_B(\omega) - \mu_{A \cap B}(\omega) \end{aligned} \tag{1.1}$$

If $\mu_{A \cap B}(\omega) = \min(\mu_A(\omega), \mu_B(\omega)) = \mu_A(\omega) \wedge \mu_B(\omega)$ then (1.1) implies $\mu_{A \cup B}(\omega) = \max(\mu_A(\omega), \mu_B(\omega)) = \mu_A(\omega) \vee \mu_B(\omega)$, whereas if $\mu_{A \cap B}(\omega) = \mu_A(\omega)\, \mu_B(\omega)$ then $\mu_{A \cup B}(\omega) = \mu_A(\omega) + \mu_B(\omega) - \mu_A(\omega)\, \mu_B(\omega)$.

To define the probability of a FS let Ω be a background space, $\mathscr{A}$ a family of FS of Ω and Q a measure on Ω which satisfies $\int_\Omega dQ = 1$. For a given $A \in \mathscr{A}$ probability of A, $P(A)$ is defined by [7]

$$P(A) = \int_\Omega \mu_A(\omega)\, dQ. \tag{1.2}$$

Since $\mu_{A^c}(\omega) = 1 - \mu_A(\omega)$,

$$P(A^c) = 1 - P(A). \tag{1.3}$$

One of the objectives of this work is to study the properties of the probability measure on FS's, $P(\ )$, the left-hand side of (1.2). It is obvious that $P(A) \geqslant 0\ \forall\ A \in \mathscr{A}$ and if we define $\mu_\Omega(\omega) \equiv 1$ then $P(\Omega) = 1$.

A FS probability calculus with the M&M algebra. The probability of the intersection of two FS's using the M&M algebra is given by [7]

$$P(A \cap B) = \int_\Omega \mu_A(\omega) \wedge \mu_B(\omega)\, dQ \tag{1.4}$$

and so

$$P(A \cup B) = \int_\Omega \mu_A(\omega) \vee \mu_B(\omega)\, dQ. \tag{1.5}$$

Some consequences are $P(A \cup A) = P(A \cap A) = P(A \cap \Omega) = P(A)$ and $P(A \cup \Omega) = 1$. If A and B are FS's then $P(A \cup B) \geqslant P(A) \vee P(B)$ and $P(A \cap B) \leqslant P(A) \wedge P(A)$. It can be shown that the traditional probability addition rule holds. That is, for FS's A and B: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. This rule may be generalized as in SPT to n FS's.

A consequence of the probability addition rule is a fundamental identity for FS:

$$P(A \cap A^c) + P(A \cup A^c) = 1 \tag{1.6a}$$

or

$$P(A \cap A^c) = P((A \cup A^c)^c). \tag{1.6b}$$

In SPT $P(A \cap A^c)$ is always zero and (1.6a) is usually written in the form $P(A \cup A^c) = P(A) + P(A^c) = 1$ since a standard set and its complement are disjoint and the probability of the union of disjoint sets is the sum of the probabilities of the sets. With FS's, on the other hand, $P(A \cap A^c)$ need not be zero. Any time $\mu_A(\omega)$ is strictly between 0 and 1 for some ω with non-zero probability $P(A \cap A^c) > 0$.

Observe that with the M&M algebra it is not necessarily true that $P(A \cap B^c) = P(A) - P(A \cap B)$. This observation will be referred to later when the notion of FS independence is taken up with the M&M algebra.

A FS probability calculus with the PR algebra: In this case

$$P(A \cap B) = \int_\Omega \mu_A(\omega)\, \mu_B(\omega)\, dQ \tag{1.7}$$

and consequently $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ in this case also.

However, $P(A \cap A) = P(A^2)$, where

$$P(A^2) = \int_\Omega \mu_A^2(\omega)\, dQ \tag{1.8}$$

and $P(A \cap A^c) = P(A) - P(A^2)$.

With the PR algebra it is true that $P(A \cap B^c) = P(A) - P(A \cap B)$ and so it will be seen that SPT independence properties carry over to FSPT if the PR algebra is used.

II. Independence and Conditional Probability

M&M algebra. A notion of FS independence was suggested by Zadeh [7], where he defined two FS's A and B to be independent if

$$P(A \cap B) = P(A)\, P(B). \tag{2.1}$$

We will now investigate the consequences of this definition and point out some of its peculiarities. Then an alternative definition of independence will be suggested that is formulated in terms of conditional probability.

With independence defined by (2.1) it is possible to show that a FS A and its complement may be independent. There are many possible examples and the following is a simple one.

	Ω	
	ω_1	ω_2
Q	1/2	1/2
μ_A	1/4	3/4
$\mu_A{}^c$	3/4	1/4

Here $P(A \cap A^c) = P(A)\, P(A^c) = \frac{1}{4}$.

If A and B are independent it may be shown that A and B^c need not be independent so that A^c and B^c need not be also. The reason is that $P(A \cap B^c) \neq P(A) - P(A \cap B)$ in general with the M&M algebra. Thus for FS's independence need not imply that of their complements.

In SPT the conditional probability of a set A given B, denoted $P(A \mid B)$, is defined by

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \tag{2.2}$$

for $P(B) > 0$. In SPT this expression is equivalent to

$$P(A \mid B) = \frac{P(A \cap B)}{P(A \cap B) + P(A^c \cap B)}, \tag{2.3}$$

since for standard sets $B = (A \cap B) \cup (A^c \cap B)$. In fact a SPT sample space partition $\{A_i\}_{i=1}^{\infty}$ may be used in the denominator of (2.2) to obtain a third equivalent form:

$$P(A \mid B) = \frac{P(A \cap B)}{\sum_{i=1}^{\infty} P(A_i \cap B)}. \tag{2.4}$$

With FS's and the M&M algebra it may be shown that these three alternatives, (2.2), (2.3) and (2.4), are not equivalent.

We adopt (2.2) as the definition of FS conditional probability, also proposed by Zadeh [7], and derive some consequences.

Since we are using the M&M algebra it is easy to see that $P(A \mid A) = 1$ and $P(A \mid \Omega) = P(A)$. If A and B are independent then $P(A \mid B) = P(A)$. This motivated our choice in the definition of conditional probability. Since the independence of A and B need not imply that of A^c and B, then it does not follow that $P(A^c \mid B) = P(A^c)$ when A and B are independent even though $P(A \mid B) = P(A)$. In addition, with the M&M algebra conditional probability need not satisfy $P(A \mid B) + P(A^c \mid B) = 1$ for $P(B) > 0$. The following equality does hold: $P(A \cup B \mid C) = P(A \mid C) + P(B \mid C) - P(A \cap B \mid C)$ for $P(C) > 0$.

The following is an example of conditional probability which seems counter-intuitive. It demonstrates that with FS's it is possible to have both $P(A) > P(A \mid B)$ and $P(A) > P(A \mid B^c)$.

	Ω		
	ω_1	ω_2	ω_3
Q	1/3	1/3	1/3
μ_A	1/4	3/4	1
μ_B	1	1/2	1/8
μ_{B^c}	0	1/2	7/8

$P(A) = 2/3$, $P(A \mid B) = 7/13$ and $P(A \mid B^c) = 5/11$ so that both $P(A) > P(A \mid B)$ and $P(A) > P(A \mid B^c)$.

It is clear that there are undesirable side effects with the assumed definitions of independence and conditional probability in the M&M case. An alternative that avoids some of these (but presents others) would be to first define conditional probability and then use this definition to define independence so that if A and B are independent so are A^c and B. To do this we define for FS's A and B, $P(B) > 0$, the conditional probability of A given B to be that of Eq. (2.3), namely,

$$P(A \mid B) = \frac{P(A \cap B)}{P(A \cap B) + P(A^c \cap B)}. \tag{2.3}$$

It is now evident that $P(A \mid B) + P(A^c \mid B) = 1$, which was not true with definition (2.2).

Next define the FS A to be *independent of the FS B* if $P(A \mid B) = P(A)$, $P(B) > 0$. In other words, $P(A)$ must satisfy

$$P(A) = \frac{P(A \cap B)}{P(A \cap B) + P(A^c \cap B)} \tag{2.5}$$

for $P(B) > 0$.

With this definition of independence it can be shown that if A is independent of B then A^c is also independent of B. However, it is also true that if A is independent of B then B *need not* be independent of A. These alternative definitions also allow the possibility to have a FS and its complement to be nontrivially independent.

PR algebra. The consequences of the independence definition (2.1) in the PR case are more in line with SPT. It may be shown that if the FS's A and B are independent so are A, B^c and A^c, B^c. It is also possible in this case to have a FS A and its complement to be independent and such that $P(A \cap A^c) > 0$.

For the PR algebra $P(A \cap A) = P(A^2)$ so the definition of conditional probability (2.2) implies $P(A \mid A) = P(A^2)/P(A) \leqslant 1$, whereas $P(A \mid A) = 1$ in the M&M case. In contrast with the M&M case since now the independence of two FS's A and B implies that of their complements it is true that $P(A \mid B) = P(A)$ implies $P(A^c \mid B) = P(A^c)$. In addition, we also $P(A \mid B) + P(A^c \mid B) = 1$, which was observed not to hold in general in the M&M case. In fact the alternative approach to conditional probability and independence is equivalent to the present one in the PR case.

III. Fuzzy Partitions and Bayes' Theorem

There is no standard way to define a fuzzy partition (FP). We will adopt a definition that is current in the FS literature. For an alternative see [1]. A sequence of FS's $\{A_i\}_{i=1}^{\infty}$ defines a FP if their characteristic functions satisfy, $\forall \omega \in \Omega$, $\sum_{i=1}^{\infty} \mu_{A_i}(\omega) = 1$. If $\{A_i\}$ is a FP and B a FS, $P(B) > 0$, then with the PR algebra $P(B) = \sum_{i=1}^{\infty} P(A_i \cap B)$ and $P(A_k \mid B) = P(A_k \cap B)/\sum_{i=1}^{\infty} P(A_i \cap B)$ the traditional Bayes' Theorem of SPT. This result does not hold in the M&M case since in that algebra in general $P(B) \neq \sum_{i=1}^{\infty} P(A_i \cap B)$, where $\{A_i\}$ is a FP.

It is evident thus far in the development of the two FS probability calculi that each calculus has advantages and disadvantages. Some SPT statements are true in both calculi. Other SPT statements are true in the M&M case that are not necessarily so in the PR case and vice versa. Neither FSPT is equivalent to SPT. Of course if all the sets are standard then all three theories agree as is usually desired of a generalization of a theory like SPT.

IV. Boole's Inequality and the Borel–Cantelli Lemmas

The main results of this section are fuzzy forms of Boole's Inequality and the Borel–Cantelli Lemmas.

Lemma 1 (Boole's Inequality). *If* $\{A_i\}$ *is a sequence of FS's then* $P(\bigcup_{i=1}^{\infty} A_i) \leqslant \sum_{i=1}^{\infty} P(A_i)$.

Proof. The proof is essentially the same as that of SPT. Let $B_k = \bigcup_{i=1}^{k} A_i$. Then

$$P(B_k) \leqslant \sum_{i=1}^{k} P(A_i) \qquad \text{(either FS algebra)}$$

thus

$$\lim_{k\to\infty} P(B_k) \leqslant \lim_{k\to\infty} \sum_{i=1}^{k} P(A_i)$$

or

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \leqslant \sum_{i=1}^{\infty} P(A_i).$$

From Boole's Inequality we obtain the Borel–Cantelli Lemma for FS's.

Lemma 2 (Borel–Cantelli). *If* $\{A_n\}$ *is a sequence of FS's and if* $\sum_{n=1}^{\infty} P(A_n) < \infty$ *then the probability that infinitely many of these occur is zero. Notationally,* $P(A_n \text{ i.o.}) = 0$, *where i.o. denotes infinitely often.*

Proof. The lemma follows from Lemma 1 as in SPT.

$$A_n \text{ i.o.} = \bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} A_k = \lim_{n\to\infty} \sup A_n.$$

By Boole's Inequality we have, for every m,

$$P(A_n \text{ i.o.}) \leqslant P\left(\bigcup_{k=m}^{\infty} A_k\right) \leqslant \sum_{k=m}^{\infty} P(A_k) \to 0$$

as $m \to \infty$ since by hypothesis $\sum_{n=1}^{\infty} P(A_n) < \infty$. Thus $P(A_n \text{ i.o.}) = 0$.

The contra-positive to Lemma 2 is that if $P(A_n \text{ i.o}) = 1$ then $\sum_{n=1}^{\infty} P(A_n) = \infty$.

To continue and prove the Borel Lemmas there would remain to be shown that if $\{A_n\}$ were a sequence of independent FS's, then $\sum_{n=1}^{\infty} P(A_n) = \infty$ implies $P(A_n \text{ i.o.}) = 1$. The proof of SPT uses the fact that independence of A and B implies that of their complements. It was pointed out that with the M&M algebra independence of A and B need not imply that of A^c and B^c. There are ways to circumvent this difficulty. The next lemma leads to one such circumvention.

LEMMA 3 (M&M algebra). *Let $\{A_n\}$ be a sequence of FS's then*

$$\text{(i)} \quad P(A_n \textit{ i.o.}) \geqslant \lim_{m\to\infty} \bigvee_{k=m}^{\infty} P(A_k)$$

or

$$\text{(ii)} \quad 1 - P(A_n \textit{ i.o.}) \leqslant \lim_{m\to\infty} \bigwedge_{k=m}^{\infty} P(A_k^c).$$

Proof.

$$\begin{aligned}
\text{(i)} \quad P(A_n \text{ i.o.}) &= P\left(\bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} A_k\right) \\
&= \lim_{m\to\infty} P\left(\bigcup_{k=m}^{\infty} A_k\right) \\
&\geqslant \lim_{m\to\infty} \bigvee_{k=m}^{\infty} P(A_k).
\end{aligned}$$

$$\begin{aligned}
\text{(ii)} \quad 1 - P(A_n \text{ i.o.}) &\leqslant 1 - \lim_{m\to\infty} \bigvee_{k=m}^{\infty} P(A_k) \\
&= \lim_{m\to\infty} \bigwedge_{k=m}^{\infty} (1 - P(A_k)) \\
&= \lim_{m\to\infty} \bigwedge_{k=m}^{\infty} P(A_k^c).
\end{aligned}$$

COROLLARY 1. *If $P(A_k) \to c$ as $k \to \infty$, $0 \leqslant c \leqslant 1$, then $P(A_n \text{ i.o.}) \geqslant c$.*

Proof. Substitute the limit c for $P(A_k)$ in (i).

Corollary 2 is one way to assure $P(A_n \text{ i.o.}) = 1$.

COROLLARY 2. *If $\lim_{m\to\infty} \bigwedge_{k=m}^{\infty} P(A_k^c) = 0$ then $P(A_n \textit{ i.o.}) = 1$.*

In summary we have shown in the M&M case that if $\{A_n\}$ is a sequence of FS's then (a) $\sum_{n=1}^{\infty} P(A_n) < \infty$ implies $P(A_n \text{ i.o.}) = 0$ and (b) if $\lim_{m\to\infty} \bigwedge_{k=m}^{\infty} P(A_k^c) = 0$ then $P(A_n \text{ i.o.}) = 1$.

A straightforward way to deduce the full Borel lemmas in the M&M case is simply to assume that the complements of the FS are independent. This is what is done in the following theorem.

THEOREM 1 (M&M algebra). *If $\{A_n\}$ is a sequence of FS's and if $\{A_n^c\}$ are independent then $P(A_n \textit{ i.o.}) = 0$ if and only if (iff) $\sum_{n=1}^{\infty} P(A_n) < \infty$ or $P(A_n \textit{ i.o.}) = 1$ iff $\sum_{n=1}^{\infty} P(A_n) = \infty$.*

Proof. All that remains to be shown is that independence of the sequence $\{A_n^c\}$ and $\sum_{n=1}^{\infty} P(A_n) = \infty$ imply $P(A_n \text{ i.o.}) = 1$. The proof is similar to the SPT argument,

$$1 - P(A_n \text{ i.o.}) = \lim_{m \to \infty} \prod_{k=m}^{\infty} (1 - P(A_k)),$$

by the assumed independence of the *complements*. The hypothesis $\sum_{n=1}^{\infty} P(A_n) = \infty$ implies the infinite product $\prod_{k=m}^{\infty} (1 - P(A_k))$ conveys to zero for all m. Thus $P(A_n \text{ i.o.}) = 1$.

With the PR algebra no difficulty is encountered in obtaining the usually stated form of the Borel lemmas.

THEOREM 1 (PR algebra). *If $\{A_n\}$ is a sequence of independent FS's then $P(A, i.o.) = 0$ iff $\sum_{n=1}^{\infty} P(A_n) < \infty$ or $P(A_n\, i.o.) = 1$ iff $\sum_{n=1}^{\infty} P(A_n) = \infty$.*

Proof. Exactly as in SPT.

V. ε-STANDARD OR ε-FUZZY SETS

We now investigate the notion of fuzzy sets that are not too fuzzy. The assumed presence of a probability measure on Ω enables a definition of "not too fuzzy" to be formalized and interesting consequences to be derived therefrom. The next section contains a generalization of the results established here.

Let P be a FS probability measure. Given such a P a pseudo-metric may be defined on the family of FS's of Ω. For FS's A and B define the pseudo-metric d by

$$d(A, B) = P(|A - B|), \tag{5.1}$$

where

$$P(|A - B|) = \int_{\Omega} |\mu_A(\omega) - \mu_B(\omega)|\, dQ.$$

It is not difficult to show that d is a pseudo-metric.

We now formalize the notion of "not too fuzzy" by defining a FS A to be *ε-standard* if there exists a standard set B such that $d(A, B) \leqslant \varepsilon$ [1]. St_ε will be the set of all ε-standard sets in the relevant context. If $0 \leqslant \varepsilon_1 \leqslant \varepsilon_2 \leqslant 1$ then of course $\mathrm{St}_{\varepsilon_1} \subset \mathrm{St}_{\varepsilon_2}$. For Fs's A and B, $A \sim_\varepsilon B$ means $d(A, B) \leqslant \varepsilon$. Thus A is ε-standard or $A \in \mathrm{St}_\varepsilon$ if there is a standard set B such that $A \sim_\varepsilon B$. For the remainder of this paper B, B_i, etc., will denote *standard* sets only.

LEMMA 4. *A is* St_ε *iff there is a standard set B such that* $P(A \cap B^c) + P(A^c \cap B) \leqslant \varepsilon$.

Proof. This lemma follows from the fact that for B standard $P(|A - B|) = P(A \cap B^c) + P(A^c \cap B)$. Note that for a standard set B

$$
\begin{aligned}
P(|A - B|) &= \int_{B \cup B^c} |\mu_A - \mu_B| \, dQ \\
&= \int_{B^c} dQ + \int_B (1 - \mu_A) \, dQ \\
&= \begin{cases} \int_\Omega \mu_A \wedge (1 - \mu_B) \, dQ + \int_\Omega \mu_B \wedge (1 - \mu_A) \, dQ & \text{(M\&M case)} \\ \int_\Omega \mu_A(1 - \mu_B) \, dQ + \int_\Omega \mu_B(1 - \mu_A) \, dQ & \text{(PR case)} \end{cases} \\
&= P(A \cap B^c) + P(A^c \cap B).
\end{aligned}
$$

Hence $P(|A - B|) \leqslant \varepsilon$ iff $P(A \cap B^c) + P(A^c \cap B) \leqslant \varepsilon$.

Lemma 4 says that if a FS A is ε-close to a standard set B, i.e., $d(A, B) \leqslant \varepsilon$, then $P(A \cap B^c) + P(A^c \cap B)$, the probability of the symmetric difference, is less than or equal to ε also.

An alternative formalization of "not so fuzzy" is as follows. A FS A could be defined to be ε-fuzzy (ε-FS) if there is a standard set B such that

$$P(A \cap B) + P(A^c \cap B^c) \geqslant 1 - \varepsilon.$$

Theorem 2 shows that this notion would be redundant.

THEOREM 2. *Let A be a FS. For all standard sets B,*

$$P(A \cap B) + P(A^c \cap B^c) \geqslant 1 - \varepsilon \textit{ iff } A \sim_\varepsilon B.$$

Proof. Necessity: Assume for some standard set B

$$P(A \cap B) + P(A^c \cap B^c) \geqslant 1 - \varepsilon.$$

It will be demonstrated that

$$P(A^c \cap B) + P(B^c \cap A) \leqslant \varepsilon.$$

Observe that

$$
\begin{aligned}
A &= A \cap \Omega = (A \cap B) \cup (A \cap B^c), \\
A^c &= A^c \cap \Omega = (A^c \cap B) \cup (A^c \cap B^c)
\end{aligned}
$$

and

$$P(A) + P(A^c) = 1.$$

Therefore

$$P(A \cap B) + P(A \cap B^c) + P(A^c \cap B) + P(A^c \cap B^c) = 1 \tag{5.2}$$

$$(\text{since } (A \cap B) \cap (A \cap B^c) = \varnothing = (A^c \cap B) \cap (A^c \cap B^c)).$$

But by assumption

$$P(A \cap B) + P(A^c \cap B^c) \geqslant 1 - \varepsilon$$

or

$$-P(A \cap B) - P(A^c \cap B^c) \leqslant \varepsilon - 1. \tag{5.3}$$

(5.2) and (5.3) imply

$$P(A \cap B^c) + P(A^c \cap B) \leqslant \varepsilon. \tag{5.4}$$

Sufficiency: Given (5.4) for some standard set B or

$$-P(A \cap B^c) - P(A^c \cap B) \geqslant -\varepsilon,$$

the latter inequality and (5.2) imply

$$P(A \cap B) + P(A^c \cap B^c) \geqslant 1 - \varepsilon.$$

The convention followed henceforth will be to call a FS A ε-fuzzy with the understanding that A is St_ε. In other words, A is ε-fuzzy iff $A \sim_\varepsilon B$, where, of course, B is a suitable standard set.

Sharp sets are 0-fuzzy. A FS with $\mu(\omega) = \frac{1}{2} \forall \omega \in \Omega$ is in a sense "most fuzzy" so it may be said a $\frac{1}{2}$-FS is the other extreme from 0-FS. It can also be shown that, without loss of generality, we may always take $0 \leqslant \varepsilon \leqslant \frac{1}{2}$. From the next lemma we obtain an example of one natural case where ε-fuzziness arises.

LEMMA 5. *$P(A \cap B^c) + P(A^c \cap B) \geqslant 1 - \varepsilon$ iff $P(A \cup B) - P(A \cap B) \leqslant \varepsilon$.*

Proof. This lemma illustrates two ways of expressing an inequality on the probability of the symmetric difference.

$$P(A \cap B) + P(A^c \cap B^c) \geqslant 1 - \varepsilon \text{ iff}$$

$$P(A \cap B) + P((A \cup B)^c) \geqslant 1 - \varepsilon \text{ iff}$$

$$P(A \cap B) + 1 - P(A \cup B) \geqslant 1 - \varepsilon \text{ iff}$$

$$P(A \cup B) - P(A \cap B) \leqslant \varepsilon.$$

An easy consequence of Lemma 5 is the following example: if B is a standard set and A a FS and either $A \subset B$ and $P(B) - P(A) \leqslant \varepsilon$ or $B \subset A$ and $P(A) - P(B) \leqslant \varepsilon$ then $A \sim_\varepsilon B$. This is because, in either case, $P(A \cup B) - P(A \cap B) \leqslant \varepsilon$.

Another convention that will be used is to say that A is nearly standard (NS) if A is ε-standard or an ε-FS for a small ε. "Small" is used here as a *fuzzy* predicate.

If A is an ε-FS so is its complement, A^c. This is a result of the definition which is symmetric in A and A^c. The goal of the next two theorems is to establish, for the M&M and PR algebras, respectively, that if $A_i \sim_{\varepsilon_i} B_i$, $i = 1, 2$, then $A_1 \cap A_2 \sim_\varepsilon B_1 \cap B_2$, where $\varepsilon = \varepsilon_1 + \varepsilon_2$. It will follow as a corollary in either case that $A_1 \cup A_2 \sim_\varepsilon B_1 \cup B_2$ too. Thus although ε-fuzziness is preserved for ε-FS complements, it is not for intersections and unions. The fuzziness of the intersection (union) of two ε-FS's is at most the sum of the fuzziness of the sets of the intersection (union). Boolean operations increase fuzziness, but this increase is a *controlled* increase.

A preliminary lemma is needed here.

LEMMA 6. *If* $f_i(\omega), g_i(\omega)\colon \Omega \to [0, 1]$ $i = 1, 2$ *then*

$$| f_i(\omega) \wedge f_2(\omega) - g_1(\omega) \wedge g_2(\omega)| \leqslant |f_1(\omega) - g_1(\omega)| + |f_2(\omega) - g_2(\omega)|.$$

Proof. We can assume without loss of generality that g_2 is the smallest. Then there are two cases:

(a) $f_2 \leqslant f_1$. In this case

$$\begin{aligned}
|f_1 \wedge f_2 - g_1 \wedge g_2| &= |f_2 - g_2| \\
&= f_2 - g_2 \\
&\leqslant |f_1 - g_1| + f_2 - g_2 \\
&= |f_1 - g_1| + |f_2 - g_2|.
\end{aligned}$$

(b) $f_1 < f_2$

$$\begin{aligned}
|f_1 \wedge f_2 - g_1 \wedge g_2| &= |f_1 - g_2| \\
&= f_1 - g_2 \\
&< f_2 - g_2 = |f_2 - g_2| \\
&\leqslant |f_1 - g_1| + |f_2 - g_2|.
\end{aligned}$$

COROLLARY. *For FS's* A_1, A_2 *and sharp sets* B_1, B_2

$$d(A_1 \cap A_2, B_1 \cap B_2) \leqslant d(A_1, B_1) + d(A_2, B_2).$$

Proof. In the statement of the lemma interpret the f_i's as the A_i FS characteristic functions and the g_i's as the B_i sharp set characteristic functions ($i = 1, 2$). Use definition (5.1) with the M&M algebra and the conclusion follows.

THEOREM 3 (M&M FS algebra). *If* $A_i \sim_{\varepsilon_i} B_i$, $i = 1, 2$, *then* $A_1 \cap A_2 \sim_\varepsilon B_1 \cap B_2$, *where* $\varepsilon = \varepsilon_1 + \varepsilon_2$.

Proof. This is an immediate consequence of the above corollary to Lemma 6.

It is now possible to show as a corollary that $A_1 \cup A_2 \sim_\varepsilon B_1 \cup B_2$, where $\varepsilon = \varepsilon_1 + \varepsilon_2$. Observe that

$$A \sim_\varepsilon B \qquad \text{iff} \quad A^c \sim_\varepsilon B^c.$$

Thus

$$A_i \sim_{\varepsilon_i} B_i \qquad \text{iff} \quad A_i^c \sim_{\varepsilon_i} B_i^c, \qquad i = 1, 2.$$

By Theorem 3

$$A_1^c \cap A_2^c \sim_\varepsilon B_1^c \cap B_2^c, \qquad \varepsilon = \varepsilon_1 + \varepsilon_2,$$

iff $(A_1^c \cap A_2^c)^c \sim_\varepsilon (B_1^c \cap B_2^c)^c$, i.e., $A_1 \cup A_2 \sim_\varepsilon B_1 \cup B_2$.

COROLLARY. $A_1 \cup A_2 \sim_\varepsilon B_1 \cup B_2$, $\varepsilon = \varepsilon_1 + \varepsilon_2$.

Remark. Theorem 3 and its corollary also hold if the B_i are FS's but we shall not need that case here.

For a proof of the consequence of Theorem 3 in the PR case a lemma similar to Lemma 6 is needed.

LEMMA 7. *Same assumptions as in Lemma* 6. *Conclusion*:

$$|f_1 f_2 - g_1 g_2| \leqslant |f_1 - g_1| + |f_2 - g_2|.$$

Proof.

$$\begin{aligned} |f_1 f_2 - g_1 g_2| &\leqslant |f_2(f_1 - g_1)| + |g_1(f_2 - g_2)| \\ &\leqslant |f_1 - g_1| + |f_2 - g_2|. \end{aligned}$$

COROLLARY (PR algebra). *For FS's* A_1 *and* A_2 *and sharp sets* B_1 *and* B_2

$$d(A_1 A_2, B_1 B_2) \leqslant d(A_1, B_1) + d(A_2, B_2).$$

Proof. As in the corollary to Lemma 6 but with the PR algebra.

Equipped with the last corollary it is possible to demonstrate that in the PR case Theorem 3 holds.

THEOREM 3 (PR algebra). *Some assumptions and conclusions as in Theorem* 3 (*M&M FS algebra*). *The proof is the same but with the PR algebra and notation and with the FSPT in use.*

The next result is a consequence of the development thus far.

THEOREM 4. *There is a standard B such that $A \sim_\varepsilon B$ only if $P(A \cap A^c) \leqslant 2\varepsilon$.*

Proof. For the standard set B

$$B \cap B^c = \varnothing.$$

Now $A \cap A^c$ need not be equal to $\varnothing$. However,

$$A \cap A^c \sim_{2\varepsilon} B \cap B^c = \varnothing,$$

i.e., $A \cap A^c \sim_{2\varepsilon} \varnothing$.

The latter implies

$$P((A \cap A^c) \cap \varnothing) + P((A \cap A^c)^c \cap \varnothing^c) \geqslant 1 - 2\varepsilon$$

so $P(((A \cap A^c) \cup \varnothing)^c) \geqslant 1 - 2\varepsilon$ and hence $1 - P((A \cap A^c) \cup \varnothing) \geqslant 1 - 2\varepsilon$ and so $P(A \cap A^c) \leqslant 2\varepsilon$.

For the converse of Theorem 4 a slightly stronger assumption is needed.

THEOREM 5. *If for a FS A, $P(A \cap A^c) \leqslant \varepsilon/2$ then there is a standard set B such that $A \sim_\varepsilon B$.*

Proof. The following applies in either case. Let $\mu_A(\omega) = x$ and $\mu_{A^c}(\omega) = 1 - x$, $0 \leqslant x \leqslant 1$. Define the standard set B as follows:

$$\begin{aligned} \mu_B(\omega) &= 1 \qquad \text{if} \quad \mu_A(\omega) \geqslant \tfrac{1}{2}, \\ &= 0, \qquad \mu_A(\omega) < \tfrac{1}{2}. \end{aligned}$$

It then follows that

$$\begin{aligned} \mu_{A \cap A^c}(\omega) &= \mu_{A^c}(\omega) = 1 - x \qquad \text{if} \quad \omega \in B, \\ &= \mu_A(\omega) = x \qquad \text{if} \quad \omega \in B^c. \end{aligned}$$

Note that

$$|\mu_A(\omega) - \mu_B(\omega)| = \mu_{A^c}(\omega), \qquad \omega \in B,$$
$$= \mu_A(\omega), \qquad \omega \in B^c.$$

Now consider the M&M case:

$$P(|A - B|) = \int_B \mu_{A^c}(\omega)\, dQ + \int_{B^c} \mu_A(\omega)\, dQ$$

$$= \int_B (\mu_{A^c}(\omega) \wedge \mu_A(\omega))\, dQ + \int_{B^c} (\mu_A(\omega) \wedge \mu_{A^c}(\omega))\, dQ$$

$$= \int_\Omega (\mu_A(\omega) \wedge \mu_{A^c}(\omega))\, dQ \leqslant \varepsilon/2 < \varepsilon.$$

PR case:

$$\varepsilon \geqslant 2 \int_\Omega \mu_A(1 - \mu_A)\, dQ$$

$$= 2 \int_B \mu_A(1 - \mu_A)\, dQ + 2 \int_{B^c} \mu_A(1 - \mu_A)\, dQ$$

$$\geqslant \int_B (1 - \mu_A)\, dQ + \int_{B^c} \mu_A\, dQ$$
$$= P(|A - B|).$$

A summary statement of the main results in this section is that the Boolean set operations intersection and union introduce more fuzziness. ε-fuzziness is preserved under complements and increases by sums with intersection or union. In the next section this observation will be used to obtain a general theorem about Boolean identities with ε-FS's.

VI. A General Theorem About Boolean Algebra Identities and ε-FS's

The notions of ε-fuzziness or near standard introduced in the previous section may be used in a natural way to obtain a general theorem about how closely these not quite standard sets satisfy Boolean identities. For example, for any standard sets B_1 and B_2 the equality

$$B_1 = (B_1 \cap B_2) \cup (B_1 \cap B_2^c) \tag{5.5}$$

holds. On the other hand, for FS's A_1 and A_2

$$A_1 \text{ need not equal } (A_1 \cap A_2) \cup (A_1 \cap A_2^c)$$

in general since $A_2 \cup A_2^c$ need not be equal to Ω.

Now suppose, however, $A_i \sim_{\varepsilon_i} B_i$, $i = 1, 2$.
Then

(i) $A_1 \sim_{\varepsilon_1} B_1$,

(ii) $A_1 \cap A_2 \sim_{\varepsilon_1+\varepsilon_2} B_1 \cap B_2$,

(iii) $A_1 \cap A_2^c \sim_{\varepsilon_1+\varepsilon_2} B_1 \cap B_2^c$ and

(iv) $(A_1 \cap A_2) \cup (A_1 \cap A_2^c) \sim_{2(\varepsilon_1+\varepsilon_2)} (B_1 \cap B_2) \cup (B_1 \cap B_2^c)$.

Putting all this together:

$$A_1 \underset{\varepsilon_1}{\sim} B_1 = (B_1 \cap B_2) \cup (B_1 \cap B_2^c) \sim_{2(\varepsilon_1+\varepsilon_2)} (A_1 \cap A_2) \cup (A_1 \cap A_2^c)$$

or

$$A_1 \sim_{(3\varepsilon_1+2\varepsilon_2)} (A_1 \cap A_2) \cup (A_1 \cap A_2^c).$$

Thus

$$d(A_1, (A_1 \cap A_2) \cup (A_1 \cap A_2^c)) \leqslant 3\varepsilon_1 + 2\varepsilon_2.$$

If $\varepsilon_1 = \varepsilon_2 = \varepsilon$ then $d \leqslant 5\varepsilon \to 0$ as $\varepsilon \to 0$.

In the sense that $A_1 \sim_\delta (A_1 \cap A_2) \cup (A_1 \cap A_2^c)$, ε-FS's almost satisfy the Boolean equality (5.5). This is an instance of a general property of ε-FS which is the essence of Theorem 6 below.

The example of Eq. (5.5) involved computing the bound $\delta = 5\varepsilon$ from $A_1 \sim_\varepsilon B_1$ and $(A_1 \cap A_2) \cup (A_1 \cap A_2^c) \sim_{4\varepsilon} (B_1 \cap B_2) \cup (B_1 \cap B_2^c)$ and the triangle inequality. The two bounds ε and 4ε for the latter formulas involving the ε-FS A_1, A_2 depend on the complexity of the formulas $\Gamma_1 = A_1$ and $\Gamma_2 = (A_1 \cup A_2) \cup (A_1 \cap A_2^c)$ in an elementary way.

To define complexity, suppose n ε-FS $A_1, \ldots, A_n$ are given and that $\Gamma(A_1, \ldots, A_n)$ is a Boolean expression. The complexity of Γ, $c(\Gamma)$ is defined to be the number of occurrences of unions and intersections plus one. That is,

$$c(\Gamma) = m + 1, \tag{5.6}$$

where m is the number of occurrences of $\cup$ and $\cap$ in Γ. Then $(A_1, \ldots, A_n) \sim_{c(\Gamma)\varepsilon} \Gamma(B_1, \ldots, B_n)$ for some standard sets $B_1, \ldots, B_n$ by assumption and induction on the results of Section V. For example, if A_1, A_2 are both ε-FS's then $c(A_1) = 1$, $c((A_1 \cap A_2) \cup (A_1 \cap A_2^c)) = 4$ and $A_1 \sim_\delta (A_1 \cap A_2) \cup$

$(A_1 \cap A_2^c)$, where $\delta = (1+4)\varepsilon$, is 5ε, the sum of the complexities times ε, as demonstrated above.

Here is another example to illustrate the present discussion. Suppose A_1 and A_2 are both 0.01-FS's. Then $A_i \sim_{0.01} B_i$ for some standard B_i, $i = 1, 2$, by assumption. From the corollary of Theorem 3, $A_1 \cup A_2 \sim_{0.02} B_1 \cup B_2$ and $(A_1 \cup A_2)^c \sim_{0.02} (B_1 \cup B_2)^c$. Observe that $c(A_1 \cup A_2) \times 0.01 = 0.02$. In addition, $(B_1 \cup B_2)^c \sim_0 B_1^c \cap B_2^c$ (this is an identity) and $A_1^c \cap A_2^c \sim_{0.02} B_1^c \cap B_2^c$. Therefore, using the triangle inequality, $(A_1 \cup A_2)^c \sim_\delta A_1^c \cap A_2^c$, where the bound $\delta = (c(A_1 \cup A_2) + c(A_1^c \cap A_2^c)) \times 0.01 = (2+2) \times 0.01 = 0.04$.

Observe that for the M&M FS algebra $(A_1 \cup A_2)^c = A_1^c \cap A_2^c$ for any FS A_1 and A_2. But for the PR algebra and ε-FS's A_1 and A_2 this equality is only "sort of" correct.

Theorem 6 generalizes the ideas discussed thus far. It shows how closely ε-fuzzy sets satisfy any Boolean equality [1].

THEOREM 6. *Let $A_1,\ldots, A_k$ be ε-fuzzy sets and suppose $\Gamma = \psi$ is an identity in Boolean set algebra involving $A_1,\ldots, A_k$. If $c(\Gamma) = m$ and $c(\psi) = l$ then*

$$\Gamma(A_1,\ldots, A_k) \sim_{(m+l)\varepsilon} \psi(A_1,\ldots, A_k).$$

The theorem says that if ε is small and m and l are not too large then for any FS's in St_ε the equality $\Gamma(A_1,\ldots, A_k) = \psi(A_1,\ldots, A_k)$ is "nearly" correct.

Proof. Let B_i $(i = 1, 2,\ldots, k)$ be standard sets such that $A_i \sim_\varepsilon B_i$. B_i exist since the A_i are in St_ε. Then

$$\Gamma(A_1,\ldots, A_k) \sim_{(m)\varepsilon} \Gamma(B_i,\ldots, B_k),$$

$$\Gamma(B_1,\ldots, B_k) \sim_0 \psi(B_1,\ldots, B_k)$$

and

$$\psi(B_1,\ldots, B_k) \sim_{(l)\varepsilon} \psi(A_1,\ldots, A_k).$$

Therefore, by the triangle inequality,

$$\Gamma(A_1,\ldots, A_k) \sim_{(m+l)\varepsilon} \psi(A_1,\ldots, A_k).$$

For fixed n, with the M&M algebra, the amount of fuzziness attainable in Theorem 6 is in fact bounded. It can be shown that there are only a finite number of Boolean expressions obtainable from n fuzzy sets $A_1, A_2 \cdots A_n$ with the M&M algebra and that an upper bound on the number of expressions is $2^{2^{2n}}$. Let c equal the maximum complexity obtainable from the n ε-fuzzy sets $A_1,\ldots,A_n$. The an upper bound on the amount of fuzziness in Theorem 6 is $2c\varepsilon$ for this fixed n. (For the PR algebra, one FS A can generate an infinite number of fuzzy sets: A, A^2, $A^3\ldots$.)

Theorem 6 may well form a semantic justification for Parikh [5], in which he argues that as tautologies in the statement calculus (which correspond to identities in Boolean algebra) based on vague premises become longer, their reliability decreases! Here, as the Boolean expressions in the identity $\Gamma = \psi$ become more complex, it is less true that $\Gamma(A_1,\ldots,A_n) = \psi(A_1,\ldots,A_n)$ for sets in St_ε. Simple expressions like $A_1 = (A_2 \cap A_2) \cup (A_1 \cap A_2^c)$ may not be too unreliable. However, added complexity involving more sets in St_ε increases the unreliability.

Acknowledgments

The results reported here are the essential parts of a doctoral dissertation completed at Boston University under the supervision of Professors Thomas A. Louis [1] and Rohit Parikh.[2] The author received much encouragement and support from them while carrying out his reasearch. In particular, the work entailed in Sections I–IV was encouraged by Tom, while that of Sections V and VI was under the guidance of Rohit. It will be seen that these sections tie together nicely.

References

1. J. F. Buoncristiani, "Probability on Fuzzy Sets," Ph. D. disssertation, Boston University, 1980.
2. J. A. Goguen, The logic of inexact concepts, *Synthese* **19** (1968)–1969), 325–373.
3. E. Hisdal, Conditional possibilities, independence and noninteraction, *Fuzzy Sets and Systems* **1** (1978), 283–297.
4. E. Hisdal, Possibilistically dependent variables and a general theory of fuzzy sets, *in* "Advances in Fuzzy Set Theory and Applications," (M. M. Gapta, R. K. Ragade, and R. R. Yager, Eds.), North-Holland, Amsterdam, in press.
5. R. Parikh, "The Problem of Vague Predicates," Research Report No. 1–77, Department of Mathematics, Boston University, February 1977.
6. L. A. Zadeh, Fuzzy sets, *Inform. and Control* **8** (1965), 338–353.
7. L. A. Zadeh, Probability measures of fuzzy events, *J. Math. Anal. Appl.* **23** (1968), 421–427.
8. L. A. Zadeh, "Fuzzy Sets as a Basis for a Theory of Possibility," Memorandum *M*77/12, Electronics Research Laboratory, University of California, Berkeley, 1977.

[1] Current address: Biostatistics Department, Harvard School of Public Health.
[2] Mathematics Department, Boston University.